Stochastic Processes and their Applications 7 (1978) 79–99.


# CONTINUITY OF A CLASS OF RANDOM TIME TRANSFORMATIONS*

Inge S. HELLAND

*Department of Mathematics, University of Bergen, Bergen, Norway*

Received 19 April 1977
Revised 9 September 1977

A random time change is defined as a map from one function space to another. The continuity of this map is investigated. Applications are made to weak limit theorems of random processes.

| | |
|---|---|
| random time transformations | Feller process |
| continuous mapping theorems | continuous state branching process |
| weak convergence | |

## 1. Introduction and definitions

By transforming the space of sample paths of random processes via a continuous transformation, it is often possible to deduce new functional limit theorems from older ones. In the first three sections of the present paper we define and investigate the continuity of one class of such transformations—corresponding to random time changes. In the last three sections we discuss various applications. Further applications will be given elsewhere.

Let $E$ be a complete, separable metric space, let $\phi$ be a continuous nonnegative function on $E$, and $B$ be a closed set in $E$ containing $\phi^{-1}(0)$. Corresponding to $\phi$ and $B$ we will introduce a map $f = f_{\phi,B}$ from the space $D_E$ of $E$-valued right continuous functions on $[0, \infty)$ with left hand limits everywhere into another space $D_E^*$ of functions on $[0, \infty)$ which is defined as follows: Let $\Delta$ be an extra point and let every $y \in D_E^*$ be a function from $[0, \infty)$ into $E \cup \Delta$ with the property that there exists a unique "lifetime" $\zeta_y \in [0, \infty]$ such that $y(t) = \Delta$ for $t \geq \zeta_y$, $y(t+) = y(t) \in E$ for $0 \leq t < \zeta_y$ and $y(t-)$ exists for $0 < t < \zeta_y$.

Now for $x \in D_E$ define

$$q_x = \inf\{t \geq 0: x(t) \in B \quad \text{or} \quad x(t-) \in B\} \tag{1.1}$$

* This paper was written while the author was visiting the University of California at Berkeley during the academic year 1976–1977. The paper is partly supported by The Norwegian Research Council for Science and the Humanities.

(where $\inf\{\emptyset\} = +\infty$) and

$$r_x = \int_0^{q_x} \mathrm{d}s/\phi(x(s)). \tag{1.2}$$

For $t < r_x$ let $\tau_x(t)$ be determined uniquely by

$$t = \int_0^{\tau_x(t)} \mathrm{d}s/\phi(x(s)). \tag{1.3}$$

Finally define the map $f = f_{\phi,B}$ by

$$f(x)(t) = \begin{cases} x(\tau_x(t)) & \text{for } 0 \leqslant t < r_x, \\ x(q_x-) & \text{for } t \geqslant r_x \text{ if } q_x < +\infty \text{ and } x(q_x-) \in B, \\ x(q_x) & \text{for } t \geqslant r_x \text{ if } q_x < +\infty \text{ and } x(q_x-) \notin B, \\ \Delta & \text{for } t \geqslant r_x \text{ if } q_x = +\infty. \end{cases} \tag{1.4}$$

Note that each of the four possibilities $q_x, r_x < +\infty$, $q_x < r_x = +\infty$, $r_x < q_x = +\infty$, $r_x = q_x = +\infty$ may occur, depending on $\phi$ and on $x$. The case $r_x < q_x = +\infty$ will be referred to later as explosion of $f(x)$. In all other cases $f(x) \in D_E$. Also note that when $q_x < +\infty$ then either $x(q_x-) \in B$, $x(q_x) \in B$ or both, since $B$ is closed and $x$ is right conttt nuous.

The object of this paper is to investigate the continuity properties of the map $f: D_E \to D_E^\ddagger$ when $D_E$ is given the Stone topology [22, 19] and $D_E^\ddagger$ is endowed with a topology that coincides with the Stone topology on its subspace $D_E$. Let $C_f = \{x \in D_E: f \text{ is continuous at } x\}$. If for any random process $X = \{X(t)\}$ one can prove that $\mathbf{P}[X \in C_f] = 1$, then one can use the continuous mapping theorem [2] to translate limit theorems of the type $X_n \to X$ (in probability, almost surely, weakly in $D_E$) into limit theorems of the type $f(X_n) \to f(X)$ [24].

As a motivation and as a source for the $X_n$-processes above, we give the following

**Theorem 1.1.** *Let $\{X(t)\}$ be a (separable) pure jump Markov process on a state space $S \subset E$ with time parameters $q(\xi)$ ($\mathbf{P}^\xi[T > t] = \exp\{-q(\xi)t\}$, where $T$ is the time of the first jump from $\xi$) and jump measure $p(\xi; A)$ ($\xi \in S$, $A \subset S - \{\xi\}$). Let $Z(t) = f(X)(t)$, where $f$ is defined by (1.1)–(1.4). Then $\{Z(t)\}$ is a Markov process on $S$ which is either pure jump or pure jump until some finite explosion time $\zeta = \zeta_Z$. The time parameters of $\{Z(t)\}$ are $q(\xi)\phi(\xi)$ for $\xi \in S - B$ and the jump measure is $p(\xi; A)$ for $\xi \in S - B$, $A \subset S - \{\xi\}$. Explosion occurs for some $\omega \in \Omega$ iff $r_{X(\omega)} < +\infty$ and $q_{X(\omega)} = +\infty$, and in this case $\zeta_{Z(\omega)} = r_{X(\omega)}$. All the states of $B \cap S$ are absorbing.*

The proof of this theorem is straightforward, either by finding the generators of the two processes, or by looking directly at the jump sizes and the time between jumps. Note that by choosing $\phi(\xi) = q(\xi)^{-1}$, we get a process $Z(t)$ whose sojourn time in each non-absorbing state $\xi$ is exponentially distributed with parameter 1.

As an application of Theorem 1.1 and the continuity theorems below, we will use well-known limit theorems for compound Poisson processes to deduce limit theorems for Markov branching processes. These constitute the continuous time analogues of the limit theorems for Galton-Watson processes given by Grimvall [12], and in particular they allow us to give necessary and sufficient conditions for the diffusion approximation of Feller/Jiřina [8, 14] to be valid.

The random time transformation studied here has also been applied to semi-stable Markov processes by Lamperti [18]. Both for this application and for the application to branching processes it is important to find conditions under which $f$ transforms Feller processes into Feller processes. This matter is discussed in Section 5 below (see also Lamperti [17]).

In all applications so far of our main result—both in the present paper and in a paper to be published on convergence to diffusion—the state space $E$ will be the real line. However, the proof of the main result is essentially the same if $E = \mathbf{R}^k$ and even if $E$ is any complete, separable metric space, so it is desirable to keep this level of generality in the formulations.

In the case $q_x = r_x = +\infty$ the definitions (1.1)–(1.4) simplify considerably. For this case the random time transformation was introduced by Volkonski [23] (see also Dynkin [6, Chapter 10]). When $q_x = r_x = +\infty$ it follows from recent results due to Whitt [24] that $f: D_E \to D_E$ is continuous at $x$. However, it seems impossible to deduce continuity of $f$ in the other cases from Whitt's theorems, and all these cases occur in the applications mentioned above.

On the notation: We will reserve the letters $x$, $y$, $z$ for functions on $E$ and denote points in $E$ by $\xi$, $\zeta$, etc. Also, for all stochastic processes we will always take the smoothest possible versions, in particular versions with sample paths in $D_E$ or $D_E^*$ whenever possible.

## 2. Main results: Continuity of the transformation

In this section we will formulate conditions under which the map $f_{\phi,B}$ defined by (1.1)–(1.4) is continuous. To this end we have to give topologies to the spaces $D_E$ and $D_E^*$. As the natural thing to do, we will endow $D_E$ with the Stone topology, i.e. $x_n \to x$ in $D_E$ iff there exists a sequence $\{\lambda_n\} \subset \Lambda$ (the space of continuous, strictly increasing functions from $[0, \infty)$ onto $[0, \infty)$) such that as $n \to \infty$

$$\sup_{0 \le t \le a} |\lambda_n(t) - t| \to 0 \tag{2.1}$$

$$\sup_{0 \le t \le a} \rho(x_n(t), x(\lambda_n(t))) \to 0 \tag{2.2}$$

for all $a > 0$. Here $\rho$ is the metric of $E$.

In the case where $E$ is non-compact and $\Delta$ is a one-point compactification of $E$, one way of giving a topology to $D_E^*$ would be to use the Stone topology on $D_{E\cup\Delta}$. It is easy to show by examples, however, that in general $f_{\Phi,B}$ will not be continuous with this topology on the image space. (Cf. examples given in Section 4.) In the sequel we will limit ourselves to the following topology of $D_E^*$, which appears to be the strongest one for which $f$ is continuous at all $x \in D_E$ such that $y = f(x)$ has finite lifetime $\zeta_y$.

**Definition 2.1.** Let the topology of $D_E^*$ be generated by the system of neighborhoods

$$U_x(c, a) = \{ y \in D_E^* : \zeta_y > c, d_c(x, y) < a \}$$

for $x \in D_E^*$, $0 < c < \zeta_x$, $0 < a < \infty$, where

$$d_c(x, y) = \inf_{\lambda \in \Lambda} \max \left\{ \sup_{0 \le t \le c} |\lambda(t) - t|, \sup_{0 \le t \le c} \rho(x(t), y(\lambda(t))) \right\}.$$

**Lemma 2.2.** *$y_n \to y$ in $D_E^*$ if and only if*

$$\liminf_{n \to \infty} \zeta_{y_n} \ge \zeta_y \tag{2.3}$$

*and there is a sequence $\{\lambda'_n\} \subset \Lambda$ such that as $n \to \infty$*

$$\sup_{0 \le t \le c} |\lambda'_n(t) - t| \to 0 \tag{2.4}$$

$$\sup_{0 \le t \le c} \rho(y_n(t), y(\lambda'_n(t))) \to 0 \tag{2.5}$$

*for all $c \in (0, \zeta_y)$.*

The proof of the lemma is immediate, and it is also easy to see (using well-known properties of the Skorokhod topology) that $D_E^*$ is a separable topological space. However, the space can not be metrized. In fact it is not even a Hausdorff space, so that limits of sequences are not unique: If $y_n \to y$, then $y_n \to y^c$ for all $c < \zeta_y$, where we define $y^c(t) = y(t)$ for $0 \le t < c$, $y^c(t) = \Delta$ for $t \ge c$. Of course in Lemma 2.2 $y_n \to y$ means that $y$ is one limit for the sequence $\{y_n\}$.

Thus the concept of convergence in $D_E^*$ is not particularly useful near $y \in D_E^*$ with $\zeta_y < \infty$. Nevertheless one may always define weak convergence $P_n \Rightarrow P$ for Borel measures $P_n$, $P$ on $D_E^*$ in the usual way by

$$\int g(y) P_n(\mathrm{d}y) \to \int g(y) P(\mathrm{d}y)$$

for every bounded continuous functional $g$ on $D_E^*$.

**Lemma 2.3.** *If $Q_n \Rightarrow Q$ in $D_E$ and $f : D_E \to D_E^*$ is continuous on some set $C \subset D_E$ with $Q(C) = 1$, then $Q_n f^{-1} \Rightarrow Q f^{-1}$ in $D_E^*$.*

**Proof.** By the Skorokhod representation theorem [3, Theorem 3.3] it is easy to see that the continuous mapping theorem is valid also when the image space of the mapping is an arbitrary topological space.

The main advantage with the above topology for $D_E^*$ is that on the subspace of non-explosive $y$'s (i.e. those with $\zeta_y = \infty$) it coincides with the Stone topology.

From now on let $\phi$ and $B$ be fixed. To study the continuity of $f = f_{\phi,B}$ we first introduce the decomposition

$$f = g \circ h, \tag{2.6}$$

where $h$ fixes the paths when points of $B$ are reached, and $g$ is the transformation $f$ restricted to the space $D_0$ of functions in $D_E$ for which all the points of $B$ are traps. Formally

$$h(x)(t) = \begin{cases} x(t) & \text{for } 0 \leqslant t < q_x, \\ x(q_x-) & \text{for } t \geqslant q_x \text{ if } x(q_x-) \in B, \\ x(q_x) & \text{for } t \geqslant q_x \text{ if } x(q_x-) \notin B, \end{cases} \tag{2.7}$$

where $q_x$ is given by (1.1). Furthermore

$$\begin{aligned} D_0 = \{x \in D_E \colon & x(t+s) = x(t) \text{ for all } s > 0 \text{ if } x(t) \in B \text{ and} \\ & x(t) = x(t-) \text{ if } x(t-) \in B\} \end{aligned} \tag{2.8}$$

and $D_0$ is endowed with the Stone topology. For $x \in D_0$ let

$$g(x)(t) = \begin{cases} x(\tau_x(t)) & \text{whenever defined}, \\ \Delta & \text{otherwise}, \end{cases} \tag{2.9}$$

where $\tau_x(t)$ is given by (1.3) for $t < r_x$, and $\tau_x(t) = q_x$ for $t \geqslant r_x$. Thus the first line in (2.9) is applicable except in the case $q_x = \infty$, $r_x < \infty$, and the definitions above are equivalent with (1.1)–(1.4).

The point is now that $g$ is a continuous transformation of $D_0$ into $D_E^*$, but $h$ may be discontinuous at some points $x \in D_E$. As a typical illustration of the latter, let $\alpha_0 \in B^c$, $\beta_0 \in \partial B$ and $\beta_n \in B^c$, $\beta_n \to \beta_0$. Define $x_n(t) = \beta_n$ for $t < 1$ and $x_n(t) = \alpha_0$ for $t \geqslant 1$ $(n \geqslant 0)$. Then $x_n \to x_0$ in $D_E$, but $h(x_n)(t) = x_n(t) = \alpha_0$ for $n \geqslant 1$, $t \geqslant 1$ and $h(x_0)(t) \equiv \beta_0$.

From now on, for $x \in D_E$ introduce the notation $b_x = h(x)(q_x)(= x(q_x \pm)$ whichever is applicable, cf. (2.7)). A sufficient condition for $h$ to be continuous at $x$ is the following.

**Condition C.** One of the following three cases holds:

C1: $q_x = +\infty$;

C2: $x(t) = b_x$ for $t \geq q_x$;

C3: $x(q_x) = b_x$ and for all $\delta > 0$ there is a $t_\delta \in [q_x, q_x + \delta)$ with $x(t_\delta) \in \operatorname{int} B$.

It is easy to see that $h$ is discontinuous at $x$ if $x(q_x) \neq x(q_x -) \in B$ or if $x(\cdot)$ leaves $B$ after $q_x$ without visiting the interior of $B$ (cf. the example above). The Condition C3 covers two completely different situations. In the first situation $x(\cdot)$ is continuous at $q_x$, but oscillates rapidly (and $B$ satisfies some regularity conditions). In the second situation $x(\cdot)$ jumps from $B^c$ to int $B$ (or to int $B$ via $\partial B$) at $q_x$. It is the first situation which is important for the applications in the last sections, and when the second situation requires special arguments, I will not always give all the details.

If $x$ satisfies Condition C, we always have $x(q_x) = b_x$ when $q_x < \infty$, and the definitions (1.1), (1.4) and (2.7) may be simplified in an obvious way.

**Proposition 2.4.** *The map $h: D_E \to D_0$ defined by* (2.7) *is continuous at any point $x \in D_E$ which satisfies Condition* C.

**Proposition 2.5.** *The map $g: D_0 \to D_E^*$ defined by* (2.9) *is continuous.*

As a consequence

**Theorem 2.6.** *The map $f: D_E \to D_E^*$ defined by* (2.6)–(2.9) (*or equivalently* (1.1)–1.4)) *is continuous at any point $x \in D_E$ which satisfies Condition* C.

The proofs of the propositions, and thereby the theorem will be deferred to the next section.

For some applications it is necessary to allow $\phi$ to take the value $+\infty$. For the purpose of the next theorem let $\phi$ be a continuous function from $E$ to $[0, +\infty]$, define $B$ as before and let $A$ be the closed set $\phi^{-1}(+\infty)$. For $x \in D_E$, let $q_x$ and $r_x$ be given by (1.1) and (1.2) respectively and define

$$\sigma_x = \inf\{t: x(t) \in A \text{ or } x(t-) \in A\}, \tag{2.10}$$

$$\rho_x = \int_0^{\sigma_x} \mathrm{d}s/\phi(x(s)). \tag{2.11}$$

When $q_x < \sigma_x$, we define $f(x)(t)$ by (1.3)–(1.4) as before. When $q_x \geq \sigma_x$ automatically $r_x \geq \rho_x$. In this case we let $\tau_x(t)$ be given by (1.3) and define $f(x)(t) = x(\tau_x(t))$ for $t < \rho_x$. For $t \geq \rho_x$ we have two possibilities, either "fix" the path: $f(x)(t) = c_x$, where $c_x = x(\sigma_x -)$ if $x(\sigma_x -) \in A$, otherwise $c_x = x(\sigma_x)$, or "kill" the path: $f(x)(t) = \Delta$ for $t \geq \rho_x$. With the first possibility we have the advantage that $f(x) \in D_E$, and convergence to $f(x)$ means convergence in the Stone topology. However, in general $f$ is only continuous at $x$ when the path is "killed" at $A$.

The following is a typical example: $E = [0, \infty)$, $\phi(\xi) = \xi^{-1}$, $A = \{0\}$; $x(t) = 1$ for $t < 1$; $x(t) = 0$ for $t \geq 1$; $x_n(t) = 1$ for $t < 1$; $x_n(t) = e^t/n$ for $t \geq 1$. Here $x_n(t) \to x(t)$ uniformly on compact intervals as $n \to \infty$. It is easy to see that $f(x) = x$ if paths are fixed, but $f(x_n)(t) \to t - 1$ as $n \to \infty$ for $t \geq 1$.

For the next theorem, note that $\sigma_x < q_x \leq \infty$ implies $\rho_x < \infty$.

**Theorem 2.7.** (a) *If $x$ satisfies Condition C and $\sigma_x > q_x$ or $\sigma_x = q_x = \rho_x = +\infty$, then $f$ is continuous at $x$.*

(b) *If $\rho_x < +\infty$, $f(x)(t) = \Delta$ for $t \geq \rho_x$ by definition of $f$ and either $\sigma_x < q_x$ or $\sigma_x = q_x = +\infty$, then $f$ is continuous at $x$.*

(c) *If $\rho_x < +\infty$, $\sigma_x < q_x$, $x(t) = c_x$ for $t \geq \sigma_x$, $f(x)(t) = c_x$ for $t \geq \rho_x$ by definition of $f$, $x_n \to x$ in $D_E$ and $\sup_{a \leq t < \infty} \rho(x_n(t), c_x) \to 0$ for some (hence all) $a > \sigma_x$, then $f(x_n) \to f(x)$ in $D_E$.*

## 3. Proofs of the main results

The proofs of Propositions 2.4 and 2.5 will be separated into several lemmas. Assume throughout that $x \in D_E$ and $\{x_n\}$ is a sequence in $D_E$ with $x_n \to x$, that is (2.1) and (2.2) hold for some fixed sequence $\{\lambda_n\}$ in $\Lambda$ and for every $a > 0$. Let $q_n$, $r_n$, $\tau_n(t)$ and $b_n$ correspond to $x_n$ as $q_x$, $r_x$, $\tau_x(t)$ and $b_x$ to $x$ (cf. (1.1), (1.2), (1.3) and the paragraph before Condition C).

**Lemma 3.1.** *Let $c \in (0, q_x)$ and $d \in (0, \infty)$. Then* $\inf_{0 \leq t \leq c} \phi(x(t)) > 0$, $\inf_{0 \leq t \leq c} \rho(x(t), B) > 0$ *and* $\sup_{0 \leq t \leq d} \phi(x(t)) < \infty$. *If $x(q_x -) \in B^c$, the first two statements hold with* $\inf_{0 \leq t \leq q_x}$ *instead of* $\inf_{0 \leq t \leq c}$.

**Proof.** By the definition of $q_x$, both $\phi(x(t-))$ and $\phi(x(t+)) = \phi(x(t))$ are positive for $t \in [0, c]$. If $\inf_{0 \leq t \leq c} \phi(x(t)) = 0$, there is a sequence $\{t_k\} \subset [0, c]$ with $\lim_k \phi(x(t_k)) = 0$. Since $\{t_k\}$ must have a convergent subsequence, and since $x(t+)$ and $x(t-)$ exist for all $t$, we get a contradiction. The next two statements follow by the same argument, and the last by an obvious modification.

**Lemma 3.2.** *Let $c \in (0, q_x)$ and $d \in (0, \infty)$. Then there is a positive integer $n_0$ such that* $\inf_{0 \leq t \leq c, n \geq n_0} \phi(x_n(t)) > 0$, $\inf_{0 \leq t \leq c, n \geq n_0} \rho(x_n(t), B) > 0$ *and* $\sup_{0 \leq t \leq d, n \geq n_0} \phi(x_n(t)) < \infty$.

**Proof.** Choose $\varepsilon > 0$ so that $c + \varepsilon < q_x$. Then $\inf_{0 \leq t \leq c + \varepsilon} \phi(x(t)) = \eta > 0$ by Lemma 3.1. Find $\delta > 0$ so small that $|\phi(\xi) - \phi(\zeta)| < \frac{1}{2}\eta$ when $\rho(\xi, \zeta) < \delta$, $\xi \in E$ and $\zeta$ belongs to the compact closure of $\{x(t)\colon 0 \leq t \leq c + \varepsilon\}$. (Compactness of the latter follows by the argument of the preceding proof.) By (2.1) and (2.2) we can find an $n_0$ so that $\sup_{0 \leq t \leq c} |\lambda_n(t) - t| < \varepsilon$ and $\sup_{0 \leq t \leq c} \rho(x_n(t), x(\lambda_n(t))) < \delta$ for $n \geq n_0$. For $n > n_0$ and

$0 \leq t \leq c$ we have $\phi(x_n(t)) \geq \phi(x(\lambda_n(t))) - |\phi(x_n(t)) - \phi(x(\lambda_n(t)))| \geq \frac{1}{2}n$. Again the rest is similar.

**Lemma 3.3.** *We always have* $\liminf_{n\to\infty} q_n \geq q_x$. *Under Condition* C3, $\lim_{n\to\infty} q_n = q_x$ *and* $\lim_{n\to\infty} b_n = b_x$.

**Proof.** By Lemma 3.2 $\liminf_{n\to\infty} q_n \geq c$ for all $c < q_x$, which proves the first part. For the rest we need the following easy consequence of (2.1), (2.2) and the triangle inequality: For all $\varepsilon > 0$ and all continuity points $t_0$ of $x(\cdot)$ there is a $\phi > 0$ so that

$$\limsup_{n\to\infty} \sup_{|t-t_0|\leq\delta} \rho(x_n(t), x(t_0)) \leq \varepsilon. \tag{3.1}$$

In particular $x_n(t_0) \to x(t_0)$ as $n \to \infty$. Now, as a consequence of C3 and right-continuity, for every $\delta > 0$ there is a continuity point $t_0$ of $x(\cdot)$ with $t_0 \in [q_x, q_x + \delta)$ and $x(t_0) \in \operatorname{int} B$. But then $x_n(t_0) \in \operatorname{int} B$ for large $n$, so $q_n \leq t_0 < q_x + \delta$ for large $n$, which shows that $\limsup_{n\to\infty} q_n \leq q_x$. If $x(q_x-) = b_x$, $x(\cdot)$ is continuous at $q_x$, and $b_n = x_n(q_n \pm) \to x(q_x) = b_x$ by (3.1). If $x(q_x-) \in B^c$, then for large $n$, $\{x_n(t)\}$ is bounded away from $B$ on $\{t: \lambda_n(t) < q_x\}$ by Lemma 3.1 and (2.2), and contained in some arbitrarily small $\varepsilon$-neighborhood of $b_x$ on $\{t: q_x \leq \lambda_n(t) < q_x + \delta\}$ for $\delta$ small enough. Then, since $q_n \to q_x$, $b_n = x_n(q_n \pm)$ must necessarily belong to this neighborhood for large $n$.

**Proof of Proposition 2.4.** We will show that (2.2) holds for all $a > 0$ when $x_n$ and $x$ are replaced by $h(x_n)$ and $h(x)$. Now

$$\rho(h(x_n)(t), h(x)(\lambda_n(t))) = \begin{cases} \rho(x_n(t), x(\lambda_n(t))) & \text{if } \lambda_n(t) < q_x, t < q_n \quad \text{(i)} \\ \rho(b_n, x(\lambda_n(t))) & \text{if } \lambda_n(t) < q_x, t \geq q_n \quad \text{(ii)} \\ \rho(x_n(t), b_x) & \text{if } \lambda_n(t) \geq q_x, t < q_n \quad \text{(iii)} \\ \rho(b_n, b_x) & \text{if } \lambda_n(t) \geq q_x, t \geq q_n \quad \text{(iv)} \end{cases} \tag{3.2}$$

In the case C1, $q_n \to \infty$ by Lemma 3.3. Then for large $n$, $q_n \geq a$ and only (3.2(i)) is possible for $0 \leq t \leq a$. In the case C2, $h(x) = x$, and (3.2) is small for $n$ large and $t$ in compact intervals satisfying (i) or (iii). Assume that $x(q_x-) = b_x$. Then by (2.2) and Condition C2, $x_n(t)$ belongs to some arbitrarily small neighborhood of $b_x$ for $n$ large, $\delta$ small and $t \in (q_x - \delta, a]$. Furthermore $q_n \geq q_x - \delta$ for large $n$. Therefore, either $q_n \geq a$, which excludes (3.2(ii)) and (3.2(iv)), or $b_n$ belongs to the same neighborhood of $b_x$, which makes (3.2) uniformly small also in cases (ii) and (iv).

Next assume that $x(\cdot)$ satisfies Condition C3 and is continuous at $q_x$. Choose $\delta$ so small that $\rho(x(s), b_x)$ is less than some given $\varepsilon$ for $|s - q_x| < \delta$. Choose $n$ so large that $|q_n - q_x| < \delta$. Then (3.2(i)) is the only possibility for $\lambda_n(t) \leq q_x - \delta$, and (3.2(iv)) is the only possibility for $\lambda_n(t) \geq q_x + \delta$. In both cases (3.2) is uniformly small. For $q_x - \delta <$

$\lambda_n(t) < q_x$ (3.2(ii)) is possible, but $\rho(b_n, x(\lambda_n(t))) \leq \rho(x(\lambda_n(t)), b_x) + \rho(b_x, b_n) < 2\varepsilon$ for large $n$. Similarly, for $q_x \leq \lambda_n(t) < q_x + \delta$ (3.2(iii)) is possible, but (3.2) is uniformly small by $\rho(x_n(t), b_x) \leq \rho(x(\lambda_n(t)), x_n(t)) + \rho(x(\lambda_n(t)), b_x)$.

It is left to discuss the case where $x(\cdot)$ is discontinuous at $q_x$, i.e. $x(q_x -) \in B^C$. Then we can use the argument in the proof of Lemma 3.3 to exclude (3.2(ii)) for large $n$. The rest of the proof goes through as before.

Next we turn to the proof of Proposition 2.5. From now on we assume that $x_n (n = 1, 2, \ldots)$ and $x$ belong to $D_0$ defined in (2.8), but otherwise we use the same notation as above. In particular (2.1) and (2.2) hold.

**Lemma 3.4.** *For every $c \in (0, q_x)$ we have*

$$\lim_{n\to\infty} \sup_{0 \leq t \leq c} \left| \int_0^t \mathrm{d}s/\phi(x(s)) - \int_0^t \mathrm{d}s/\phi(x_n(s)) \right| = 0.$$

**Proof.**

$$\left| \int_0^t \mathrm{d}s/\phi(x(s)) - \int_0^t \mathrm{d}s/\phi(x_n(s)) \right| \leq \int_0^c \left| 1/\phi(x(s)) - 1/\phi(x_n(s)) \right| \mathrm{d}s \to 0$$

since $x_n(s) \to x(s)$ for all continuity points of $x(\cdot)$, $\phi$ is continuous, and we can use dominated convergence by Lemma 3.2.

**Lemma 3.5.** *For every $d \in (0, r_x)$ we have*

$$\lim_{n\to\infty} \sup_{0 \leq t \leq d} |\tau_n(t) - \tau_x(t)| = 0.$$

**Proof.** By Lemma 3.4, $\tau_n^{-1}(t)$ converges uniformly to $\tau_x^{-1}(t)$ on $[0, c]$, where $c = \tau_x^{-1}(d)$. Furthermore the slope of the continuous function $\tau_x^{-1}(t)$ (i.e. $1/\phi(x(t))$) is bounded away from 0 and $\infty$ by Lemma 3.1. The uniform convergence of $\tau_n(t)$ to $\tau_x(t)$ follows by a simple geometric consideration.

From now on let $y(t) \equiv g(x)(t)$ and $y_n(t) \equiv g(x_n)(t)$ with $g$ defined by (2.9). Choose $c \in (0, \xi_y)$.

**Lemma 3.6.** *There exists a positive number $a = a(x, c)$ and a natural number $n_0$ such that*

$$\tau_x([0, c]) \cup \bigcup_{n=n_0}^{\infty} \tau_n([0, c]) \cup \bigcup_{n=n_0}^{\infty} \lambda_n(\tau_n([0, c])) \subseteq [0, a]. \tag{3.3}$$

*If $c < r_x$, we can choose $a < q_x$.*

**Proof.** By the monotonicity of $\tau_n(\cdot)$ and by (2.1) it is enough to show that $\{\tau_n(c): n \geq n_0\}$ is strictly bounded by $a$ for $n_0$ large enough. If $c < r_x$ this follows immediately from Lemma 3.5. Since $\zeta_y = r_x$ if $q_x = \infty$, we have only left the case $q_x < \infty$, $r_x < \infty$, $c \geq r_x$. For this case let $a > q_x + \delta$ for some $\delta > 0$. Then $\sup_{q_x+\delta \leq t \leq a} \rho(x_n(t), b_x) \to 0$ and we can find $n_0$ such that $\sup_{q_x+\delta \leq t \leq a} \phi(x_n(t)) \leq \phi(b_x) + \delta$ for $n \geq n_0$. Therefore

$$\tau_n^{-1}(t) = \int_0^t ds/\phi(x_n(s)) \geq \int_{q_x+\delta}^t ds/\phi(x_n(s)) \geq (t - q_x - \delta)/(\phi(b_x) + \delta)$$

when $q_x + \delta \leq t \leq a$ and $n \geq n_0$. If we have $a > q_x + \delta + c(\phi(b_x) + \delta)$, then $\tau_n^{-1}(a) > c$ and therefore $\tau_n(c) < a$. Note that if $\phi(b_x) = 0$ (as when $B = \{\xi: \phi(\xi) = 0\}$), then we can choose as $a$ any number larger than $q_x$.

**Proof of Proposition 2.5.** We have to prove (2.3)–(2.5) for some sequence $\{\lambda'_n\}$, and every $c < \zeta_y$. The relation (2.3) follows easily from the first part of Lemma 3.3 and from Lemma 3.4, since $\zeta_y = r_x$ if $q_x = \infty$ and $\zeta_y = \infty$ otherwise, and correspondingly $\zeta_{y_n} = r_n$ or $+\infty$ (cf. (1.4)).

From now on we fix $c \in (0, \zeta_y)$ and choose $a$ as in Lemma 3.6. For this $a$ we will use (2.1) and (2.2). We first have to construct the sequence $\{\lambda'_n\}$. (It will be seen that the construction is independent of $a$ and $c$).

Let $\{u_n\}$ be any monotone sequence of positive numbers with $u_n \uparrow r_x$, $u_n < r_n$ and

$$\lambda_n(\tau_n(u_n)) < q_x \quad \text{for every } n. \tag{3.4}$$

It is possible to find such a sequence since $\tau_x(r_x) = q_x$ and by (2.1) and Lemma 3.5 $\lambda_n(\tau_n(u)) \to \tau_x(u)$ as $n \to \infty$ for every $u < r_x$. Define for $0 \leq u \leq u_n$

$$\lambda'_n(u) = \int_0^{\lambda_n(\tau_n(u))} ds/\phi(x(s)), \tag{3.5}$$

which is possible by (3.4). For $u > u_n$ we will extend $\lambda'_n$ linearly (some modification will be done later for the case where $x(\cdot)$ is discontinuous at $q_x$).

$$\lambda'_n(u) = \lambda'_n(u_n) + u - u_n \quad \text{for } u > u_n. \tag{3.6}$$

Then every $\lambda'_n \in \Lambda$, and by (3.5)

$$\tau_x(\lambda'_n(u)) = \lambda_n(\tau_n(u)) \quad \text{for } u \in [0, u_n]. \tag{3.7}$$

Therefore, again for $u \in [0, u_n]$

$$\begin{aligned} \rho(y_n(u), y(\lambda'_n(u))) &= \rho[x_n(\tau_n(u)), x(\tau_x(\lambda'_n(u)))] \\ &= \rho[x_n(\tau_n(u)), x(\lambda_n(\tau_n(u)))] \end{aligned} \tag{3.8}$$

which is uniformly small for $n$ large and $u \in [0, u_n] \cap [0, c]$ by (2.2) and (3.3).

Furthermore

$$|\lambda_n'(u)-u| = \left|\int_0^{\lambda_n(\tau_n(u))} \mathrm{d}s/\phi(x(s)) - \int_0^{\tau_x(u)} \mathrm{d}s/\phi(x(s))\right| \\ = \left|\int_{\tau_x(u)}^{\lambda_n(\tau_n(u))} \mathrm{d}s/\phi(x(s))\right| \tag{3.9}$$

for $u \leq u_n$.

The easiest case is when $r_x = \infty$. Then $u_n > c$ for $n$ large enough, so that (2.5) holds by (3.8). Also $\lambda_n(\tau_n(u)) \to \tau_x(u)$ uniformly in $u \in [0, c]$ by Lemma 3.5 and (2.1), and (2.4) follows from (3.3), (3.9) and Lemma 3.1. The explosion case $r_x < \infty$, $q_x = \infty$ is similar. Here $c < \zeta_y = r_x$ and $u_n > c$ for $n$ large enough.

Next consider the case where $r_x < \infty$, $q_x < \infty$ and $x(q_x -) = b_x \in B$. Again there is no problem when $c < r_x$, since we then only need the definitions for $u \leq u_n$. For $u > u_n$ we have

$$\rho(y_n(u), y(\lambda_n'(u))) = \rho[x_n(\tau_n(u)), x(\tau_x(\lambda_n'(u_n)+u-u_n))] \\ \leq \rho[x_n(\tau_n(u)), b_x] + \rho[x(\tau_x(\lambda_n'(u_n)+u-u_n)), b_x]. \tag{3.10}$$

Since $x(t) = b_x$ for $t \geq q_x$ and is continuous at $q_x$ and since for $u \geq u_n$

$$\tau_x(\lambda_n'(u_n)+u-u_n) \geq \tau_x(\lambda_n'(u_n)) = \lambda_n(\tau_n(u_n)) \to q_x, \tag{3.11}$$

the second term in (3.10) tends to zero uniformly in $u$. The first term also tends to zero uniformly for $u$ in compact intervals, since for every $\varepsilon > 0$, $\sup_{q_x-\delta \leq t \leq a} \rho(x_n(t), b_x) < \varepsilon$ when $\delta$ is small enough and $n$ large enough (cf. (2.2)). This, together with (3.8) proves (2.5). To prove (2.4), it is enough to show that the righthand side of (3.9) is uniformly small for $n$ large and $u \leq u_n$, since $\lambda(u) - u = \lambda(u_n) - u_n$ for $u > u_n$. To this end, choose first $\varepsilon > 0$ and then $\delta > 0$ so small that

$$\int_{q_x-\delta}^{q_x} \mathrm{d}s/\phi(x(s)) < \varepsilon. \tag{3.12}$$

This is possible by (1.2) since $r_x < \infty$. For some fixed $u_0 < r_x$ and $n$ large enough both $\tau_x(u)$ and $\lambda_n(\tau_n(u))$ are larger than $q_x - \delta$ for $u \geq u_0$. Then $|\lambda_n'(u) - u| < \varepsilon$ for $u_0 \leq u \leq u_n$ by (3.9) and (3.12), and for $u \leq u_0$, $\lambda_n'(u) \to u$ uniformly by (2.1), Lemma 3.5 and Lemma 3.4.

It is only left to discuss the case where $q_x$ and $r_x$ are finite and $x(\cdot)$ is discontinuous at $q_x$. Then the argument around (3.11) breaks down, and we have to be more careful when defining $\lambda_n'(u)$ for $u > u_n$. We use the inequality

$$\rho(y_n(u), y(\lambda_n'(u))) = \rho[x_n(\tau_n(u)), x(\tau_x(\lambda_n'(u)))] \\ \leq \rho[x_n(\tau_n(u)), x(\lambda_n(\tau_n(u)))] + \rho[x(\lambda_n(\tau_n(u))), x(\tau_x(\lambda_n'(u)))]. \tag{3.13}$$

The first term on the righthand side is again uniformly small for $n$ large and $u$ in compact intervals. The second term is zero for $u \leq u_n$ and depends on $\lambda_n'$ for $u > u_n$.

Let $s_n$ be defined by $\lambda_n(\tau_n(s_n)) = q_x$ (or $s_n = r_n$ if this is not possible) and let $\lambda'_n(s_n) = r_x$. By (3.4) we see that $s_n > u_n$ and by (3.5), $\lambda'_n(s_n) > \lambda'_n(u_n)$. Let $\lambda'_n(u)$ be defined by linear interpolation for $u_n < u < s_n$ and by $\lambda'_n(u) = \lambda'_n(s_n) + u - s_n$ for $u > s_n$. Then $x(\lambda_n(\tau_n(u)))$ and $x(\tau_x(\lambda'_n(u)))$ are both equal to $b_x$ when $u \geq s_n$ and both are near $x(q_x -)$ when $u_n \leq u < s_n$. To prove (2.4) we have to show that $s_n \to r_x$ when $n \to \infty$. The details are left to the reader.

**On the proof of Theorem 2.7.** In case (a) the proof of Theorem 2.6 goes through as before. In case (b) and (c) we can modify this proof, replacing $B$, $q_x$, $q_n$ etc. by $A$, $\sigma_x$, $\sigma_n$ etc. The extra hypotheses allow us to avoid the use of Lemma 3.6, which do not hold here.

## 4. Limit theorems for random processes

Again let $E$ be a complete, separable metric space. The purely topological Theorem 2.6 is the basis for the following theorem of probabilistic nature. Typical applications are found by combining it with Theorem 1.1.

**Theorem 4.1.** *Let $\{X_n(t): n \geq 1\}$ be a sequence of random processes with state spaces $S_n \subseteq E$ and with paths in $D_E$. Suppose that $\{X_n(t)\}$ converges weakly in $D_E$ to some limit process $\{X(t)\}$ with state space $E$, and suppose that the paths of $\{X(t)\}$ satisfy Condition* C *almost surely for some given nonnegative continuous function $\phi$ on $E$ and a closed set $B \supset \{\xi \in E: \phi(\xi) = 0\}$. Then the sequence of processes $\{f_{\phi,B}(X_n)(t)\}$ with $f_{\phi,B}$ given by* (1.1)–(1.4) *converges weakly in $D_E^*$ to some limit process which has the representation $\{f_{\phi,B}(X)(t)\}$. If* $\mathbf{P}[r_X < q_X = \infty] = 0$, *the convergence may be strengthened to weak convergence in $D_E$ (i.e. with the Stone topology).*

**Proof.** This is a direct application of Theorem 2.6 and the version of the continuous mapping theorem given in Lemma 2.3. By the Skorokhod representation theorem $\{X_n(t)\}$ and $\{X(t)\}$ may be represented on the same probability space such that $X_n(\cdot) \to X(\cdot)$ almost surely in $D_E$. For every path $X(\cdot)$ with $q_X < \infty$ or $r_X = \infty$, convergence to $f_{\phi,B}(X)(\cdot)$ in $D_E^*$ is equivalent to convergence in the Stone topology.

To use this theorem one has to verify that one of the Conditions C1–C3 is satisfied. A typical example for which one can easily prove C3, is the following: Let $X(\cdot)$ be Brownian motion in $E = \mathbf{R}^k$ and let $B$ be a closed set in $\mathbf{R}^k$. If we assume that all points of $B$ are regular for $\text{int}(B)$ in the sense discussed e.g. in Dynkin and Yushkevich [7, Chapter II, §§ 5–6], then C3 is immediate. In fact $B$ can be any closed set such that the Brownian motion do not start at an irregular boundary point of $B$. It is a classical result from potential theory that the irregular boundary points form a polar set, i.e., one that is a.s. not hit by $X(\cdot)$.

It is desirable to give some further discussion of the explosion case. When the probability of explosion is nonzero, Theorem 4.1 is not particularly useful, since weak convergence in $D_E^*$ does not even imply convergence in finite-dimensional distributions. Consider the following deterministic example: Let $y_n(t)=1/(1-t)$ for $0\leq t<1-1/n$, $y_n(t)=1$ for $t\geq 1-1/n$ and let $y(t)=1/(1-t)$ for $0\leq t<1$, $y(t)=\infty$ (or $\Delta$) for $t\geq 1$. Then $y_n(\cdot)\to y(\cdot)$ in $D_E^*$, but $y_n(t)\nrightarrow y(t)$ for $t>1$. On the other hand, if we try to give a stronger topology to the image space, the continuity of the transformation $f$ ceases to hold. In the example above, $y_n=f(x_n)$ and $y=f(x)$, where $f$ is given by (1.1)–(1.4) with $E=\mathbf{R}^1$, $\phi(\xi)=\xi^+$ and $B=(-\infty, 0]$ and where $x_n(t)=\mathrm{e}^t$ for $0\leq t<\log n$, $x_n(t)=1$ for $t\geq\log n$, $x(t)\equiv\mathrm{e}^t$. Obviously $x_n\to x$ in the Stone topology, but it is impossible to obtain $f(x_n)\to f(x)$ in any topology which concerns $f(x)(t)$ for $t$ larger than explosion time.

However, by imposing relatively weak additional conditions on the sequence $\{X_n\}$ in Theorem 4.1, we can still prove convergence in finite-dimensional distributions of $\{f_{\phi,B}(X_n)\}$.

**Condition D.** For the case where $\mathbf{P}[r_X<q_X=\infty]>0$ and $\{X_n(t)\}$ converges weakly in $D_E$ to $\{X(t)\}$, suppose that

$$\lim_{t,n\to\infty}\mathbf{P}[X_n(t)\in K]=0 \tag{4.1}$$

for all compact sets $K\subset E$.

**Theorem 4.2.** *If Condition* D *is satisfied in addition to the conditions of Theorem* 4.1 *and if $\{f_{\phi,B}(X)(t)\}$ has no fixed discontinuities, then as $n\to\infty$ the finite-dimensional distributions of $\{f_{\phi,B}(X_n)(t)\}$ converge to the (possibly defective) finite-dimensional distributions of $\{f_{\phi,B}(X)(t)\}$.*

**Proof.** To avoid notational mess, we will only prove convergence of 1-dimensional distributions. Let $Y_n(t)=f_{\phi,B}(X_n)(t)$, $Y(t)=f_{\phi,B}(X)(t)$ and let $\psi$ be a continuous real function on $E$ with compact support $K$. We have to prove

$$\lim_{n\to\infty}E\psi(Y_n(u))=E\psi(Y(u)) \tag{4.2}$$

for every $u>0$. The conclusion is independent of the representation of the processes. Since $D_E$ is a complete, separable metric space, we can by the Skorokhod representation theorem define $X_n(\omega,t)$ and $X(\omega,t)$ on one and the same probability space $(\Omega,\mathscr{A},\mathbf{P})$ in such a way that $X_n(\omega,\cdot)\to X(\omega,\cdot)$ in $D_E$ for every $\omega\in\Omega$. If $\omega$ is such that $q_X(\omega)=\infty$, then $Y_n(\cdot,\omega)$ converges to $Y(\cdot,\omega)$ in the Skorokhod topology on $D_E[0,a]$ for every $a<r_X(\omega)$ by Theorem 2.6. Therefore, for any fixed $u>0$,

$$\lim_{n\to\infty}E(\psi(Y_n(u));u<r_X,q_X=\infty)=E(\psi(Y(u));u<r_X,q_X=\infty).$$

Similarly

$$\lim_{n\to\infty} E(\psi(Y_n(u)); q_X < \infty) = E(\psi(Y(u)); q_X < \infty).$$

If $u \geqslant r_X(\omega)$ and $q_X(\omega) = \infty$, then by definition $Y(u, \omega) = \Delta$ and $\psi(Y(u, \omega)) = 0$ (see (1.4)). By Lemma 3.3 $q_n(\omega) \to \infty$. (Here $q_n$, $r_n$, $\tau_n$ correspond to $X_n$ as $q_X$, $r_X$, $\tau_X$ to $X$.) By Lemma 3.5 and the monotonicity of $\tau_n$, for every $u_0 < r_X(\omega)$,

$$\liminf_{n\to\infty} \tau_n(u, \omega) \geqslant \liminf_{n\to\infty} \tau_n(u_0, \omega) = \tau_X(u_0, \omega),$$

and since $\tau_X(r_X) = q_X = \infty$ therefore

$$\lim_{n\to\infty} \tau_n(u, \omega) = \infty.$$

Thus either $Y_n(u, \omega) = \Delta$ or $Y_n(u, \omega) = X_n(\tau_n(u, \omega), \omega)$ with $\tau_n(u, \omega) \to \infty$ as $n \to \infty$. By (4.1), and since a.s. convergence of $\tau_n(u, \cdot)$ implies convergence in distribution

$$\lim_{n\to\infty} \mathbf{P}[Y_n(u) \in K, u \geqslant r_X, q_X = \infty] = 0.$$

Finally, by Lebesgue's dominated convergence theorem for convergence in probability

$$\lim_{n\to\infty} E(\psi(Y_n(u)); u \geqslant r_X, q_X = \infty) = E(\psi(Y(u)); u \geqslant r_X, q_X = \infty) = 0.$$

Thus (4.2) holds.

**Remark 4.3.** If $\phi(\xi) = +\infty$ on some set $A$, if paths are killed when hitting $A$, if (4.1) holds for all compacts $K \subset E - A$, and if $\{f(X)(t)\}$ has no fixed discontinuities, then again $\{f(X_n)(t)\}$ converges to $\{f(X)(t)\}$ in finite dimensional distributions as $\{X_n\}$ converges weakly in $D_E$ to $X$. (We assume again that $X$ satisfies Condition C a.s.) The above proof works with the same translations as in the proof of Theorem 2.7.

## 5. Markov processes and the Feller property

In this section let $X = \{X(t)\}$ be a homogeneous, strong Markov process with state space $(E, \mathcal{B})$, where $E$ is a complete, separable metric space and $\mathcal{B}$ is its $\sigma$-algebra of Borel sets. Let the transition function of $X$ be $\mathbf{P}_t(\xi, U)$ for $t \geqslant 0$, $\xi \in E$, $U \in \mathcal{B}$. We will assume that $X$ is uniformly stochastically continuous in the sense that for all $\varepsilon > 0$

$$\lim_{t\to 0} \sup_{\xi\in E} \mathbf{P}_t(\xi, V_\varepsilon(\xi)) = 0, \tag{5.1}$$

where $V_\varepsilon(\xi) = \{\zeta \in E : \rho(\xi, \zeta) \geq \varepsilon\}$. Then (e.g. by [21], Theorem 2.3) we may without loss of generality assume that the sample paths of $X$ are right-continuous with lefthand limits everywhere, i.e. belong to $D_E$. We assume that $X$ is a conservative process, although this requirement may be relaxed.

Let $\phi$ be a continuous function from $E$ into $[0, \infty]$, $B$ be a closed set containing $\phi^{-1}(0)$ and $A = \phi^{-1}(+\infty)$. Let $f = f_{\phi,B}$ be given by (1.1)–(1.4) in case $A$ is empty and in one of the two ways discussed in connection with Theorem 2.7 when $A$ is nonempty. Let $Y = \{Y(t)\}$, where $Y(t) = f(X)(t)$. Using the arguments in Volkonskii [23], suitably modified to allow for finite lifetime, we can show that $Y$ is again a homogeneous strong Markov process. In general the process is non-conservative, and the sample paths belong to $D_E^*$.

In addition we will assume that $X$ is a Feller process, that is for every real, bounded, continuous function $\psi$ on $E$ and every $t > 0$

$$E^\xi \psi(X(t)) = \int_E \psi(\zeta) \mathbf{P}_t(\xi, \mathrm{d}\xi) \tag{5.2}$$

is a continuous function of $\xi$. In [17] Lamperti found conditions under which $Y$ is also a Feller process. The following theorem extends some of Lamperti's results. Further extensions are possible; in particular (5.1) may be relaxed somewhat.

**Theorem 5.1.** *Let $N$ be a subset of $E$ such that $A \subset N \subset E$, and for all $\xi \in E - N$ the following hold* a.s. $[\mathbf{P}^\xi]$: *The paths of $X$ satisfy Condition* C *and the paths of $Y$ do not hit $N$. Furthermore* $\mathbf{P}^\xi[r_X < q_X = \infty] = 0$ *for all $\xi \in E - N$. Then $Y$ is a Feller process on $E - N$.*

**Proof.** Choose $\xi \in E - N$, $\xi_n \in E - N$, $\xi_n \to \xi$. Let $X = \{X(t)\}$ be started at $\xi$ and let $X_n = \{X_n(t)\}$ be copies of the same process started at $\xi_n$. The proof will be complete if we can show that for every such choice and for every fixed $t > 0$

$$f(X_n)(t) \Rightarrow f(X)(t). \tag{5.3}$$

(As usual $\Rightarrow$ denotes weak convergence). By the Feller property, $X_n \to X$ in finite-dimensional distributions. Now (5.1) is sufficient for the tightness-condition in [21, Theorem 3.2] to apply, so that $X_n \Rightarrow X$ (in $D_E$).

Let $\bar{\mathbf{P}}$ be the measure on $D_E$ induced by $X$ and fit $t > 0$. By [5, Theorem 6.7] and by (5.1) $X$ is quasileftcontinuous, from which we easily deduce that the mapping $s \to f(x)(s)$ is continuous at $s = t$ a.s. $[\bar{\mathbf{P}}]$. If $f(x)(\cdot)$ is continuous at $t$, the projection from $D_E$ to $E$ given by $y \to y(t)$ is continuous at $y = f(x)$. Thus (5.3) will follow from the continuous mapping theorem if we can show that $f$ is continuous a.s. $[\bar{\mathbf{P}}]$, and by our assumptions this follows directly from Theorem 2.7(a) (with $N$ replacing $A$).

**Remark 5.2.** We have nowhere assumed the state space $E$ to be locally compact. That it is a complete, separable metric space is assumed (at least implicitly) in Lamperti's paper too.

**Remark 5.3.** Lamperti's condition given by [17, eq. (4.3)] is somewhat weaker than our Condition C3. This is a reflection of the fact that our main results are formulated in a way which also covers non-Markov processes.

**Remark 5.4.** Assume that $E$ is locally compact, noncompact and let $C_0$ be the space of continuous functions that vanish at infinity. For the explosion case and for the case where $Y$ is killed when hitting $A$, we can easily give conditions under which the semigroup corresponding to $Y$ leaves $C_0$ invariant. Theorem 2.7(b), Theorem 4.2 and Remark 4.3 are applied.

**Remark 5.5.** Similarly we can deduce analogues of [17, Theorem 4] from our Theorem 2.7(c).

## 6. Convergence of a sequence of Markov branching processes

In [15] Lamperti defined the continuous state space branching (C.B.) processes as the class of Markov processes on $[0, \infty)$ (with Borel $\sigma$-algebra) whose transition function $P_t(\xi, \cdot)$ is non-trivial, jointly measurable in $\xi$ and $t$ and satisfies the branching property

$$\mathbf{P}_t(\xi+\zeta, \cdot) = \mathbf{P}_t(\xi, \cdot) * \mathbf{P}_t(\zeta, \cdot) \tag{6.1}$$

for all $t, \xi, \zeta \geq 0$. (Here $*$ denotes convolution).

Lamperti showed that the class of C.B. processes coincides with the class of possible limit processes for a sequence of the type

$$Y_n(t) = Z_{[nt]}^{(n)}/b_n, \tag{6.2}$$

where $\{Z_k^{(n)}\}_n$ is a sequence of Galton–Watson processes started at $Z_0^{(n)} = c_n$ and it is assumed that $c_n \to +\infty$. In [12] Grimvall showed that the sequence (6.2) converges to a particular limit process $\{Y(t)\}$ (necessarily of C.B. type) if and only if the row sums of a certain triangular array of independent random variables converge to a corresponding infinitely divisible distribution. To achieve this result Grimvall used rather elaborate manipulations of generating functions.

The purpose of this section is to derive analogues of Grimvall's results for a sequence of continuous time Markov branching processes. For the case where the offspring distributions have finite means with upper limit less than or equal to one, and where the $\lambda_n$ below do not depend on $n$, the result in [12] may be derived from Theorem 6.1 below and vice versa, using another result of Grimvall [11], stating that the limiting behavior of a sequence of Bellman–Harris processes depends on the life-length distribution only through its mean.

Instead of (6.2), let us write

$$Y_n(t) = Z_{nt}^{(n)}/b_n. \tag{6.3}$$

where for each $n$, $\{Z_t^{(n)}\}$ is a Markov branching process with offspring distribution $\{p_k^{(n)}: k=0,2,3,\dots\}$ and exponentially distributed lifetime with mean $1/\lambda_n$. Let $Z_0^{(n)}=c_n$ and assume that $c_n\to\infty$ as $n\to\infty$. It is easy to see that [15, Theorem 1] carries over to this situation, so that if the sequence defined by (6.3) converges in finite-dimensional distributions (or even in one-dimensional distributions) to some non-trivial limit process $Y=\{Y(t)\}$, then $Y$ must be a C.B. process.

The point is now that

$$Y_n=f(X_n), \tag{6.4}$$

where $f$ is given by (1.1)–(1.4) with $E=\mathbf{R}^1$, $\phi(\xi)=\xi^+=\max(\xi,0)$ and $B=(-\infty,0]$. Here $X_n$ is a compound Poisson process started at $X_n(0)=c_n/b_n$ and

$$E[\exp\{-u(X_n(s+t)-X_n(s))\}]=\exp\{\chi_n t(\psi_n(u)-1)\}, \tag{6.5}$$

where $\chi_n=n\lambda_n b_n$ and

$$\psi_n(u)=\sum_{k=0}^{\infty} p_k^{(n)}\exp\{-ub_n^{-1}(k-1)\}. \tag{6.6}$$

This representation follows directly from Theorem 1.1 and well-known properties of the Markov branching processes. A reverse representation theorem—a special case of which is given in [1, Theorem 1, p. 126]—also follows from our Theorem 1.1.

For formulation of the next theorem let $\mu_n$ be the measure on $\mathbf{R}^1$ with characteristic function $\psi_n(u)$, i.e. with mass $p_k^{(n)}$ concentrated on the points $(k-1)/b_n$ $(k=0,2,3,\dots)$. Let $\mu$ be an infinitely divisible probability measure on $\mathbf{R}^1$ with positive spectral part, i.e.,

$$\int_{-\infty}^{\infty}\exp\{-u\xi\}\mu(\mathrm{d}\xi)=\exp\{-\psi(u)\} \tag{6.7}$$

where

$$\psi(u)=\alpha u-\tfrac{1}{2}\beta u^2+\int_{(0,1]}(1-e^{-u\theta}-u\theta)\nu(\mathrm{d}\theta)+\int_{(1,\infty)}(1-e^{-u\theta})\nu(\mathrm{d}\theta) \tag{6.8}$$

for $u$ real; here $\nu(1,\infty)$ and $\int_{(0,1]}\theta^2\nu(\mathrm{d}\theta)$ are finite, $\alpha$ is real and $\beta\geq 0$.

For the sequence of processes $\{Y_n(t)\}$ given in (6.3) to converge to a limit, it is necessary that $c_n/b_n$ converges to some positive limit $\gamma$. (Put $t=0$.) This will be assumed in the following. Let $X=\{X(t)\}$ be the spectrally positive Lévy process started at $X(0)=\gamma$ and with exponent $\psi(u)$ given by (6.8), i.e.

$$E(\exp\{-u(X(s+t)-X(s))\})=\exp\{-t\psi(u)\}. \tag{6.9}$$

Let $Y=\{Y(t)\}$ be given by $Y=f(X)$, with the same $f$ as in (6.4). Note that explosion may occur. For $E=\mathbf{R}^1$ we write $D=D_E$ and $D^*=D_E^*$.

**Theorem 6.1.** *The following statements are equivalent*:

(a) $(\mu_n)^{*k_n} \Rightarrow \mu$, *where* $k_n = [n\lambda_n b_n]$.

(b) $X_n = \{X_n(t)\}$ *converges in finite-dimensional distributions to* $X = \{X(t)\}$.

(c) $X_n \Rightarrow X$ *in* $D$.

(d) $Y_n = \{Y_n(t)\}$ *converges in finite-dimensional distributions to* $Y = \{Y(t)\}$.

(e) $Y_n$ *converges in finite-dimensional distributions to* $Y$, *and if* $Y$ *is non-explosive then also* $Y_n \Rightarrow Y$ *in* $D$.

**Corollary 6.2.** *Necessary and sufficient conditions for the sequence* $\{Y_n\}$ *in* (6.3) *to converge weakly in the Stone topology to a diffusion process* $Y$ *started at* $Y(0) = \gamma$ *and with generator*

$$\tfrac{1}{2}\beta\xi\frac{\partial^2}{\partial\xi^2} + \alpha\xi\frac{\partial}{\partial\xi}$$

*are the following*:

(i) $c_n/b_n \to \gamma$.

(ii) *For all* $\varepsilon > 0$, $n\lambda_n b_n \sum_{k \geq \varepsilon b_n} p_k^{(n)} \to 0$.

(iii) *For all* $\varepsilon > 0$, $n\lambda_n \sum_{|k-1| < \varepsilon b_n} (k-1)p_k^{(n)} \to \alpha$.

(iv) *For all* $\varepsilon > 0$, $n\lambda_n b_n^{-1} \sum_{|k-1| < \varepsilon b_n} (k-1)^2 p_k^{(n)} \to \beta$.

**Proof.** This follows from Theorem 6.1(a) and standard theorems on convergence of row sums of triangular arrays to a normal limit (e.g. [10, Theorem 2, p. 128]). Note that if the limits in (iii) and (iv) hold for one $\varepsilon > 0$, they hold for all $\varepsilon > 0$ by (ii).

To prove Theorem 6.1 we need the following

**Lemma 6.3.** *Let* $\{X_n(t)\}$ *be a sequence of Lévy processes converging in finite-dimensional distributions to a nondegenerate Lévy process* $\{X_0(t)\}$. *Then for all* $a > 0$

$$\lim_{n,t\to\infty} \mathbf{P}[|X_n(t)| \leq a] = 0. \tag{6.10}$$

**Proof.** Assume without loss of generality that $X_n(0) = 0$, and let $G_{n,t}(\xi) = \mathbf{P}[X_n(t) \leq \xi]$; $\exp\{-t\phi_n(u)\} = E\exp\{iuX_n(t)\}$ for $n \geq 0$. Then

$$\begin{aligned}\mathbf{P}[|X_n(t)| \leq a] &= \int_{-a}^{a} \mathrm{d}G_{n,t}(\xi) \\ &\leq A^{-1}\int_{-a}^{a} \xi^{-2}(1 - \cos(\xi/a))\,\mathrm{d}G_{n,t}(\xi) \\ &\leq (2A)^{-1}\int_{s=0}^{a^{-1}}\int_{u=-s}^{s} \exp\{-t\phi_n(u)\}\,\mathrm{d}u\,\mathrm{d}s,\end{aligned}$$

where $A = \inf_{|\xi| \le a} \xi^{-2}(1 - \cos(\xi/a)) > 0$. The last inequality follows from

$$\int_{-s}^{s} \exp\{-t\phi_n(u)\}\,du = \int_{-s}^{s}\int_{-\infty}^{\infty} e^{iu\xi}\,dG_{n,t}(\xi)\,du = \int_{-\infty}^{\infty} 2\xi^{-1}\sin(s\xi)\,dG_{n,t}(\xi).$$

From well-known properties of characteristic functions [9, p. 501], there is a number $\lambda > 0$ such that

$$|E e^{iuX_0(1)}| = \exp\{-\operatorname{Re}\phi_0(u)\} < 1 \quad \text{for } 0 < |u| < \lambda.$$

Without loss of generality we may suppose that $a > \lambda^{-1}$, i.e. $a^{-1} < \lambda$. Then for all $\delta < 0$ there is a positive integer $n_0$ and a $c \in (0, 1)$ such that

$$\exp\{-\operatorname{Re}\phi_n(u)\} \le c \quad \text{for } n \ge n_0,\ \delta \le |u| \le a^{-1}$$

(since $\phi_n \to \phi_0$ uniformly on compact sets). For $n \ge n_0$ we have

$$\mathbf{P}[|X_n(t)| \le a] \le (2A)^{-1}\int_0^{a^{-1}} (\delta + (s-\delta)c^t)\,ds \le (2Aa)^{-1}(\delta + c^t/2a),$$

from which (6.10) follows, since $\delta$ is arbitrary.

**Proof of Theorem 6.1.** (a) *and* (b) *are equivalent*: It is easy to see that $\psi_n(u)^{k_n} \to e^{-\psi(u)}$ if and only if $k_n(\psi_n(u)-1) \to -\psi(u)$, i.e. if and only if $\chi_n(\psi_n(u)-1) \to -\psi(u)$, since $\chi_n/k_n \to 1$. (Instead of Laplace transforms, we may also use the corresponding characteristic functions and refer to the elementary lemma in Feller [9, p. 555].) Thus (a) is equivalent to convergence in one-dimensional distribution in (b), which implies finite-dimentional convergence, since all processes have stationary, independent increments.

(b) *implies* (c): By Skorokhod [20, Theorem 2.7], convergence in finite-dimensional distributions of Lévy processes implies weak convergence in $D[0, 1]$, similarly in $D[0, r]$ for every $r > 0$, and therefore in $D = D[0, \infty)$.

(c) *implies* (e): We will first show that $Y_n \Rightarrow Y$ in $D^*$. This is immediate from Theorem 4.1 if we can show that the process $X$ satisfies Condition C a.s. Two cases can occur. Either $X$ is a subordinator (has a.s. non-decreasing paths) in which case $q_X = +\infty$ a.s., or $q_X < +\infty$ with positive probability, $X(q_X) = 0$ a.s. when $q_X < +\infty$, and $q_X$ is a stopping time for the process. Thus $X(q_X + t)$, conditioned on $q_X < +\infty$, has the same distribution as $X(t) - X(0)$. In this latter case, when $X$ is not a subordinator, it either has negative drift $\alpha - \int_{(0,1]} \theta\nu(d\theta)$ (see (6.8)) or its paths are a.s. of unbounded variation in every compact interval ($\int_{(0,1]} \theta\nu(d\theta) = +\infty$ or $\beta > 0$ or both). Then it is easy to see that a.s. there are arbitrarily small $t$'s with $X(t) - X(0) < 0$. Thus when $q_X < +\infty$, there are $t$'s arbitrarily close to $q_X$ with $t > q_X$ and $X(t) < 0$, and hence either Condition C1 or Condition C3 is satisfied a.s.

To prove convergence of finite-dimensional distributions in the explosion case, we use Theorem 4.2. That Condition D is satisfied is immediate from Lemma 6.3.

(e) *implies* (d): Trivial.

(d) *implies* (b): The Kolmogorov forward equation for the process $Y_n$ defined in (6.3) gives

$$\frac{\partial}{\partial t}E(\mathrm{e}^{-uY_n(t)})=n\lambda_n b_n(\psi_n(u)-1)E(Y_n(t)\,\mathrm{e}^{-uY_n(t)}).$$

Thus for all $u>0$

$$\begin{aligned}\lim_{n\to\infty}\chi_n(\psi_n(u)-1)&=\lim_{n\to\infty}\{E(\mathrm{e}^{-uY_n(t)}-\mathrm{e}^{-uY_n(0)})\}\left\{\int_0^t E(Y_n(s)\,\mathrm{e}^{-uY_n(s)})\,\mathrm{d}s\right\}^{-1}\\&=\{E(\mathrm{e}^{-uY(t)}-\mathrm{e}^{-uY(0)})\}\left\{\int_0^t E(Y(s)\,\mathrm{e}^{-uY(s)})\,\mathrm{d}s\right\}^{-1}\\&=\phi(u),\end{aligned}\tag{6.11}$$

say. We note that $\phi(u)\to 0$ as $u\to 0$, even in the explosion case, when $\mathbf{P}[Y(t)=+\infty]>0$. Thus by (6.5), $\{X_n(t)\}$ converges in distribution to a nondegenerate limit. Necessarily the limit process must have stationary, independent increments, and by the first part of the proof it is identified with $\{X(t)\}$.

Once Theorem 6.1 is established, we can re-derive several previous results about C.B.-processes from it. For instance, it is immediate that every limit process that occur in (d) and (e) has a representation of the form $Y=f(X)$, where $X$ is a spectrally positive Lévy process and $f$ is as in (6.4). This is just the representation for C.B.-processes given by Lamperti in [16]. Also, we can derive from (6.11) the differential equation (cf. [4, (1.10)] or [11, Theorem 3.1(b)]) connecting the Laplace transform of the distribution of $Y$ with the function $\psi$ in (6.7)–(6.9).

As another application of Theorem 2.6, we mention that it allows us to transform central limit type theorems into general theorems about convergence of random processes to diffusion processes [13]. The condition for convergence found in this way can be shown to be minimal in a certain sense, and the theorems apply to any sequence of processes with paths in $D$.

## References


[1] K.B. Athreya and P.E. Ney, Branching Processes (Springer, Berlin, 1972).

[2] P. Billingsley, Convergence of Probability Measures (Wiley, New York, 1968).

[3] P. Billingsley, Weak Convergence of Measures: Applications in Probability (SIAM, Philadelphia, PA, 1971).

[4] N.H. Bingham, Continuous branching processes and spectral positivity, Stochastic Processes Appl. 4 (1976) 217–242.

[5] E.B. Dynkin, Die Grundlagen der Theorie der Markoffschen Prozesse (Springer, Berlin, 1961).

[6] E.B. Dynkin, Markov Processes, Vol. I (Springer, Berlin, 1965).

[7] E.B. Dynkin and A.A. Yushkevich, Markov Processes, Theorems and Problems (Plenum Press, New York, 1969).

[8] W. Feller, Diffusion processes in genetics, Proc. 2nd Berkeley Symp. Math. Statist. Probability (Univ. of Calif. Press, Berkeley, CA, 1951) 227–246.

[9] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. II (Wiley, New York, 1971) 2nd ed.
[10] B.V. Gnedenko and A.N. Kolmogorov, Limit Distributions for Sums of Independent Random Variables (Addison-Wesley, Reading, MA, 1968).
[11] A. Grimvall, Approximation of Bellman–Harris branching processes, Report LiH-MAR-R-73-6, Dept. of Mathematics, Linköping Univ., Linköping (1973).
[12] A. Grimvall, On the convergence of sequences of branching processes, Ann. Probability 2 (1974) 1027–1045.
[13] I.S. Helland, Minimal conditions for weak convergence to a diffusion process on the line, submitted to Ann. Probability.
[14] M. Jiřina, On Feller's branching diffusion processes, Časopis Pěst. Mat. 94 (1969) 84–90.
[15] J. Lamperti, The limit of a sequence of branching processes, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 7 (1967) 271–288.
[16] J. Lamperti, Continuous state branching processes, Bull. Am. Math. Soc. 73 (1967) 382–386.
[17] J. Lamperti, On random time substitutions and the Feller property, in: J. Chover, ed., Markov Processes and Potential Theory (Wiley, New York, 1967).
[18] J. Lamperti, Semi-stable Markov processes I, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 22 (1972) 205–225.
[19] T. Lindvall, Weak convergence of probability measures and random functions in the function space $D[0, \infty)$, J. Appl. Probability 10 (1973) 109–121.
[20] A.V. Skorokhod, Limit theorems for stochastic processes with independent increments, Theor. Probability Appl. 2 (1957) 138–171.
[21] A.V. Skorokhod, Limit theorems for Markov processes, Theor. Probability Appl. 3 (1958) 202–246.
[22] C. Stone, Weak convergence of stochastic processes defined on semi-infinite time intervals, Proc. Am. Math. Soc. 14 (1963) 694–696.
[23] V.A. Volkonski, Random substitution of time in strong Markov processes, Theor. Probability Appl. 3 (1958) 310–326.
[24] W. Whitt, Continuity of several functions on the function space $D$, to appear.